ON BOTH SIDES OF THE QUALITY BIAS IN PRICE INDEXES

BART HOBIJN
Federal Reserve Bank of New York*

Version 3, June 2003

Abstract

It is often argued that price indexes do not fully capture the quality improvements of new goods in the market. Because of this shortcoming, price indexes are perceived to overestimate the actual price increases that occur. In this paper, I argue that the quality bias in price indexes is just as likely to be upward as it is to be downward. I show how both the sign and the magnitude of the quality bias in the most commonly applied price index methods is determined by the cross-sectional variation of prices per quality unit across the product models sold in the market.

I do so by introducing a model of a market that includes monopolistically competing suppliers of the various product models and a representative consumer with CES (Constant Elasticity of Substitution) preferences. I simulate the model using actual observed CPU prices and find that a large part of the price declines measured for CPUs turns out to be due to a downward quality bias rather than to actual price declines.

Keywords: Price index theory, hedonic price methods, quality bias, monopolistic competition.
JEL-codes: C43, D11, D43, L13.

* Correspondence: Bart Hobijn, Domestic Research Function, Federal Reserve Bank of New York, 33 Liberty Street, 3rd floor, New York, NY 10045. E-mail: bart.hobijn@ny.frb.org. The views expressed in this paper are those of the author and do not necessarily reflect those of the Federal Reserve Bank of New York or those of the Federal Reserve System as a whole. I would like to thank Alexis Antoniades and Ashwin Vasan for their excellent research assistance. I would like to thank seminar participants at the European University Institute as well as Omar Licandro and Kevin Stiroh for their comments and suggestions.

1. Introduction

There is a widespread consensus among economists that price index methods tend to overestimate actual inflation in markets where there is a rapid turnover of goods due to technological progress. The Boskin (1996) commission made this point with respect to the U.S. Consumer Price Index, while Gordon (1990) used hedonic price indexes to correct for this bias in equipment price indexes.

There is, however, also a small number of studies that challenge this conventional wisdom. Studies by by Triplett (1972,2002), Feenstra (1995), as well as Hobijn (2001) have each made the point that quality adjustments in price index methods might actually lead to an understatement of inflation.

This paper follows up on the above papers by introducing a parsimonious theoretical model that can generate both a positive as well as a negative quality bias in the most commonly applied price indexes. The value added of this approach is that it allows for the study of the factors that determine the sign and magnitude of the quality bias in a stylized framework. This contrasts strongly with the methodology that is traditionally applied in the price index literature.

A large part of the literature on price indexes compares various price indexes calculated for the same dataset. This is for example the approach of Aizcorbe and Jackman (1993), Manser and McDonald (1988), and Braithwait (1980) when assessing the magnitude of substitution bias as well as of Aizcorbe, Corrado, and Doms (2000) and Silver and Heravi (2002) in the comparison of hedonic and matched model price indexes.

Such an approach allows us to consider the sensitivity of price indexes to the choice of method applied. It does not, however, enable us to make any normative statements about which index method is 'better' than another. Such normative statements on price indexes are all based on the extensive theoretical price index literature, which focuses on properties like idealness, exactness, and superlativity of price index formula.

It turns out that the theoretical results derived in this paper contradict some of the properties of price indexes that are presumed in this applied strand of the literature. Three results stand out in particular.

The most important is that the theoretical model in this paper confirms the claims by Triplett (1972,2002), Feenstra (1995), and Hobijn (2001) that the quality bias in price indexes is not by definition upward. Moreover, the sign and magnitude of the bias turn out *not to depend* on the overall level of inflation. Instead they depend on the cross-sectional behavior of prices per quality unit across models sold in the market during the same period.

Secondly, the existence and sign of this bias does not depend on the specific price index formula applied. I show how the application of the most popular price index formulas, like Laspeyres, Paasche, Geometric Mean, Fisher Ideal, and Tornqvist, all lead to a bias in the same direction.

Finally, hedonic price indexes suffer from the same quality bias as matched model indexes. Hence, the theoretical results here seem to disagree with the presumption that hedonic price indexes do a better job at correcting prices for quality improvements, as made in, among many, Pakes (2002) and Hulten (2002).

The particular theoretical model that I use for my analysis in this paper is that of a market with a representative consumer with CES preferences over a set of models sold. This setup is very similar to Dixit and Stiglitz (1977) and Hornstein (1993). The main difference is that the market that I consider has a countably finite number of models and suppliers. The advantage of this choice of model is that price index theory for CES preferences is extremely well developed. Sato (1976) derived the ideal exact price index for CES preferences when the same models are sold in both the base- and measurement periods. Feenstra (1994) extended Sato's index to an exact matched model index that can be used when the universes of models sold in both periods do not coincide.

The methodology used in this paper is closely related to the Monte Carlo methodology in econometrics. In this sense, I follow Lloyd (1975) who also used simulation methods to quantify certain properties of price indexes. In Lloyd's (1975) study the focus was on the substitution bias in price indexes. Here the focus is on their quality bias. I take Lloyd's (1975) approach one step further by not generating my own data but instead basing my simulation on observed prices of CPU's for PCs. Hence, my simulation results do not only illustrate the existence of the bias but also its empirical relevance.

The structure of the paper is as follows. In the next section I introduce the form of the CES preferences that I consider in the rest of the paper and derive the theoretical price level that price indexes are meant to measure. In Section 3 I illustrate graphically how conventional price index methods might yield a downward quality bias for these preferences. This graphical description is essentially an informal version of the results that are derived in the context of the theoretical model. I introduce this theoretical model in Section 4. I consider its demand and supply side and show how its Pure Strategy Nash equilibrium exists and is unique. In Section 5 I then proceed by deriving some general results for the sign of the quality bias in matched model and hedonic price indexes calculated for a specific parameterization of the model. In Section 6 I present the results of a simulation of the model that is based on actual data on prices and benchmark ratings of CPUs. I illustrate how this simulation confirms the results shown in Sections 3 through 5 and show why these results are empirically relevant. Section 7 concludes.

2. CES-preferences and the theoretical price level

The aim of this paper is to be able to make normative statements about price index methods and to say which ones perform better, in certain situations, than others. In order to make these normative statements we need to define what it is we would like our price index methods to measure. Since Konüs (1939) the main focus of price index theory has been on constructing a cost-of-living index (COLI). The aim of a COLI is to track the (percentage) changes in the minimum expenditures required to reach a certain base-level of utility over time[1].

The minimum amount of expenditures that is necessary to reach a certain utility level crucially depends on the underlying preferences of the consumer. Hence, the theoretical price level that price index methods are after depends on the preferences of the consumer. In reality, a market consists of a spectrum of consumers with different preferences. It turns out that it is not always possible, in such cases, to specify the theoretical price level because aggregate demand does not always behave as if it is generated by a well-behaved aggregate utility representation.

[1] I will focus on consumer price indexes throughout this paper. The theory presented in this paper is also applicable to producer price indexes, which are aimed at tracking the minimum cost required to obtain a base quantity of output over time.

The focus of this paper is not on the conditions for the existence of an aggregate utility representation for aggregate demand. What I will do is simply use one of the best developed aggregate utility representations for which it has been proven that it can be interpreted as the aggregate utility function of a market with a continuum of heterogeneous agents. This aggregate utility representation is Constant Elasticity of Substitution (CES) preferences. Anderson, de Palma, and Thisse (1993) introduced the microfoundations of CES preferences and showed how they can be interpreted as the aggregate utility representation of a market consisting of a continuum of heterogeneous agents.

Let $X_{i,t}$ be the quantity consumed of good i at time t, where I will assume that good $i=0$ is the numeraire good. C_t is the universe of goods sold at time t. I will assume that aggregate demand in the market in the theoretical model behaves as if resulting from the utility maximizing decision of a representative consumer with the utility function

$$U_t = \left(\sum_{i \in C_t \backslash \{0\}} \left(a_i X_{i,t} \right)^{\frac{1}{1+\lambda}} \right)^{1+\lambda} X_{0,t}^{\alpha} \text{ where } \lambda>0 \text{ and } 0<\alpha<1 \tag{1}$$

This is a relatively standard CES utility function, where $\sigma \equiv (1+\lambda)/\lambda$ is the constant elasticity of substitution. The only non-standard features of (1) are that the quantities for goods $i \in C_t \backslash \{0\}$ are multiplied by a quality parameter a_i and that the numeraire good, $i=0$, is included.

Let $p_{i,t}$ be the price of a unit of good i. Since $i=0$ is the numeraire good, I will assume that $p_{0,t}=1$ for all t. In the rest of this paper, I will focus on the construction of a price index for the set of goods, which I will call models in the future, that are contained in the CES part of (1). That is, my focus is on the measurement of the price level of the set of models $i \in C^*_t$ where $C^*_t = C_t \backslash \{0\}$.

We are thus confronted with two sets of goods, i.e. the numeraire good and the models for which we would like to measure an aggregate price level. Diewert (2001) shows that, because the preferences in (1) are separable between $X_{0,t}$ and the other goods, the aggregate price level for the models $i \in C^*_t$ is well defined. In particular, aggregate demand for the models $i \in C^*_t$ will be as if it was generated by the representative agent maximizing the amount of utility obtained from these models for the expenditures solely on these models. This implies that the theoretical price aggregate for the set of models $i \in C^*_t$ is the CES price aggregate as applied in, among many, Dixit and Stiglitz (1977), Hornstein (1993), and Feenstra (1994). This aggregate, the value at time t of which I will denote by P_t^T, reads

$$P_t^T = \left[\sum_{i \in C_t^*} \left(p_{it} / a_i \right)^{-1/\lambda} \right]^{-\lambda} \tag{2}$$

It is a CES aggregate of the prices per quality unit for all models that are traded in the market. This price aggregate represents the money cost of a unit of utility obtained from the consumption of the competing varieties in the set of models C_t^*. This money cost does not depend on the base-level of utility because the preferences are homothetic.

The aim of price index methods is to construct an index that approximates, up to a constant, the path of P_t^T. In particular, the index methods are meant to estimate the period by period percentage change in P_t^T.

Throughout this paper, I will focus on the percentage change in P_t^T between periods $t=0$ and $t=1$. I will refer to the percentage change in P_t^T between those two periods as the theoretical inflation rate and will denote it as

$$\pi^T = \frac{P_1^T - P_0^T}{P_0^T} \tag{3}$$

It represents the percentage change of the money cost of a unit of utility between periods 0 and 1.

If one would know all the preference parameters in (1) then it would not be difficult to calculate the theoretical inflation rate in (3). In practice, however, the preference parameters are not observed. That is, we do not exactly know the elasticity of substition, i.e. $\sigma \equiv (1+\lambda)/\lambda$. Neither do we know the quality embodied in each unit sold for each model, i.e. a_i. In fact, when we apply price index methods we do not even know by what preference representation aggregate demand is generated. There are basically two lines of thought here, which I will both pursue in this paper.

The first line assumes that aggregate preferences belong to a certain class and then uses this restriction to obtain an estimate of (3). For the CES preferences the index that exactly measures the theoretical inflation rate is the one derived by Sato (1976). The details of this index are described in Table 1. Sato's index is valid under the assumption that the universes of models sold in both periods are the same, such that $C_0^* = C_1^*$. It is a proper price index in the sense that it only depends on observables, namely expenditure shares and prices.

The requirement of coinciding sets of models being sold in both periods renders the Sato (1976) index inapplicable at many lower levels of aggregation. Many markets have a high rate of product turnover, as illustrated in Aizcorbe, Corrado, and Doms (2000) for the market for Intel CPU units and in Silver and Heravi (2002) for the market for laundry machines. Hence, it is thus essential to develop price index methods that allow for dynamic universes of models that change over time, i.e. $C_0^* \neq C_1^*$. Feenstra (1994) extends Sato's result to a quasi-index that is exact for CES preferences with non-overlapping universes of models. Feenstra's is a quasi-index because it depends on the unobserved elasticity of substitution, which has to be estimated to implement the index. It is described in Table 1 and I will discuss its intuition in more detail later on.

Price index theory is thus very well developed for CES preferences. We know the form of the exact indexes both when the universe of models is static as well as when it is dynamic. The problem is that in many practical cases it is a big leap to assume that demand is generated by aggregate CES preferences. This brings us to the second line of thought. This line is to construct price indexes that do not exactly measure (3) but instead reasonably approximate it for a very broad class of preferences.

This is the approach most commonly chosen for the calculation of aggregate statistics. Classical price index theory, among others Konüs (1939), Frisch (1936) and Fisher (1922), yielded many important results for the case in which the universe of models is static. Konüs (1939) introduced the concepts of a cost-of-living index and substitution bias in price indexes that price a fixed basket of goods. Frisch (1936) showed how Konüs's substitution bias result implied that for homothetic preferences the change in the true cost of living is bounded from above by the Laspeyres index and from below by the Paasche index. Fisher (1922) showed how the geometric mean of the Laspeyres and Paasche indexes constitutes an ideal index in the sense that both the price and quantity indexes have the same functional form.

A large part of the literature has focused on the question which price index formula approximates (3) in the 'most reasonable' way. Examples of studies along this line are Fisher (1922), Diewert (1976), as well as Lloyd (1975), Braithwait (1980), Manser and McDonald (1988), and Aizcorbe and Jackman (1993).

A much smaller part of the literature has focused on the construction of 'reasonable' approximations to (3) in case of dynamic universes of models. The problem when the universes of models are dynamic is that the prices of new goods are not observed in the first period, while the prices of obsolete goods are not observed in the second period. It is thus not possible to measure the percentage change in the prices between both periods for new and obsolete goods.

Two approaches are generally considered when dealing with this problem. The first, known as matched model indexes, makes specific assumptions about the relative price per quality unit of the new models versus the old models. These assumptions are such that they imply that the change in the overall price level can be estimated solely as a function of the price changes of the models that are sold in both periods, i.e. that are matched. Triplett (2002) contains an overview of the different matched model methods and the possible biases that they induce.

The second, known as hedonic price indexes, uses a regression model that relates the price of a model in a certain period to its characteristics to impute the unobserved prices for the new and obsolete models. This imputation completes the set of prices needed to apply conventional price index methods developed for overlapping universes of models. After the price imputation of the missing price observations, indexes are then constructed using conventional price index methods.

3. A graphical illustration of the main argument

The conventional wisdom is that the introduction and obsolescence of goods in a market would cause standard price index methods to overstate the actual inflation rate. The Boskin (1996) commission report as well as its recent reassessment by Lebow and Rudd (2001) both contain extensive descriptions of this conventional wisdom. There are three main reasons why this is argued to be the case. The first reason, designated *quality bias* by the Boskin (1996) commission, is that current price indexes do not properly capture the quality improvements embodied in new (or improved) models. By underestimating these quality improvements, price indexes will attribute too much of changes in expenditures to changes in prices rather than to changes in quantities. The second reason, designated *product bias* by the Boskin (1996) commission, argues that prices of new goods tend to drop faster than those of established models. Because new goods and models are only included in the sample of goods used to calculate the price index with a certain delay, the initial price drops early in the product cycle are not captured by current price indexes. Finally, there is the *substitution bias*. This bias is due to some price indexes, including the CPI and most price indexes calculated in Europe, being fixed weighted price indexes which do not capture the increases in welfare from consumers being able to substitute new goods for goods that they were previously consuming.

In the rest of this paper I will mainly focus on the *quality bias* and ignore issues related to the latter two sources of bias. In general it is hard to argue against statistical agencies including new models and goods more timely in their samples and reducing the potential sources of product bias. Furthermore, the issue of substitution bias is currently being addressed, at least for the U.S. CPI, by the joint publication of a fixed weighted as well as a chain weighted price index. The latter is meant to account for the substitution bias. See Bureau of Labor Statistics (2002) for a detailed description.

The main point of this paper is that the *quality bias* in price indexes is not solely a source of upward bias. Instead, the quality bias induced by most commonly applied price index methods can be both upward as well as downward. Before I illustrate this in a formal mathematical economic framework, I first describe the main intuition of the argument graphically in this section. The graphical description in this section is based on Figure 1 through Figure 3.

The top panel of Figure 1 depicts two hypothetical price schedules, for $t=0$ and $t=1$, of a set of models that differ according to their quality levels, a_i. I will assume that a_i is not directly observed. Therefore, the researcher observes the price of each model, i.e. p_{it}, but does not know its relative position on the x-axis. As explained in the previous section, what is important for the price level associated with the CES preferences that I consider is not the actual price levels, p_{it}, but the price per unit of quality, p_{it}/a_i, for each model. Panel (b) of Figure 1 depicts the associated schedule of prices per quality unit. Panel (c) contains the same price per quality unit schedule and adds some of the notation that I will use in the rest of this paper.

Just like in the previous section C^*_t denotes the set of models sold in period t, while P_t^T denotes the theoretical price level at time t. Note that I have chosen to draw the example such that $P^T{}_0<P^T{}_1$. That is, in the graphical example the actual price level increases between periods $t=0$ and $t=1$, such that there is positive inflation. In each period the set of models sold, i.e. C^*_t, consists of a group of models that are not sold in the other period, i.e. the set A_t-B_t, as well as a group of models that are 'matched' in the sense that they are sold in both periods, i.e. the set B_t-D_t.

What I will now illustrate is that, even though the theoretical price inflation is positive for these hypothetical price schedules, most commonly applied price index methods will tend to measure negative inflation instead. That is, in this graphical example standard price index methods will tend to *underestimate* actual inflation rather than overestimate it, as the consensus view suggests. I will illustrate this for both matched model as well as hedonic price indexes.

There are several ways in which matched model indexes are calculated. They each make different identifying assumptions about the relative price per quality unit of the obsolete and new models in the market.

The first method, often referred to as 'direct comparison', assumes that the obsolete and new models can be directly compared in the sense that they embody the same levels of quality. Because this method assumes that there are no quality improvements between the old and new vintages of models, this method is never applied in markets with rapid product turnover due to technological progress, like those for computers and other electronic products for example. Because I will focus on markets with quality improvements in the products sold, I will disregard this method in the rest of this paper.

The second method, known as 'link-to-show-no-price-change', assumes that the price per quality unit is the same for the obsolete and new models. In this case, the relative price of the obsolete and new models is assumed to be fully attributable to quality improvements. Aizcorbe (2001) uses this assumption for example to identify the parts in semiconductor price changes attributable to quality changes and price changes respectively. Note that, as Triplett (2002) describes in more detail, this method *overestimates* inflation only when the price per quality unit of the new models is lower than that of the obsolete models. In that case the method overestimates the price per quality unit for the new models and thus will overestimate inflation. The reverse is true in our graphical example here. In the example the price per quality unit is higher for the new models than for the old models. Consequently, the method will overestimate quality changes and *underestimate* the actual level of inflation.

The final matched model method that is frequently applied is the "Implicit Price – Implicit Quantity"[2] method (IP-IQ). This method is based on the identifying assumption that the overall price change equals the price change in the set of matched models. When one makes this assumption, the price levels of the unmatched models are not needed to measure inflation. Hence, in this case the unmatched models are ignored, i.e. "deleted", and standard price index methods are applied to the set of models that is sold in both periods, i.e. the matched models.

Figure 2 illustrates the application of the IP-IQ method in our graphical example. The set of matched models in the example is the intersection of C^*_0 and C^*_1. Consequently, the IP-IQ method will compare the B_0-D_0 part of the period *0* price schedule with the B_1-D_1 part of that of period *1*. For all models in this range the prices are falling. The IP-IQ method will thus, incorrectly, find a drop in the overall price level. The simplest way to see why the IP-IQ method underestimates inflation in this case is to compare the relative prices of the deleted sections A_0-B_0 and A_1-B_1 with the matched parts of the price schedules.

For the deleted part A_0-B_0 in period *0* we obtain that the prices per quality unit are lower than the prices per quality unit on the matched part of the schedule, B_0-D_0. Consequently, the deletion of the below average prices on the A_0-B_0 part of the price schedule will lead to an inferred price level in period *0* that is higher than the actual level. Similarly, when the above average prices of part A_1-B_1 are deleted in period *1*, the prices of the matched models, i.e. B_1-D_1, reflect a price level that is lower than the actual price level. That is, because the price per quality unit is increasing in quality and the worst models become obsolete while the new models are of the highest quality, the price level in period *0* is overestimated while the price level in period *1* is underestimated. The combination of these two measurement errors leads to an unambiguous downward bias in the measured inflation rate, independent of which price index formula is applied.

One final thing is worth noting about this argument. That is that the bias incurred due to the application of the IP-IQ method does not depend on the overall inflation rate. Instead, it completely depends on the cross-sectional behavior of the prices per quality unit as a function of the quality units embodied in the models sold in the market. I will prove this in a more formal example later on. This result contrasts sharply with the argument in Triplett (2002) who argues that "The errors produced by the IP-IQ method are symmetric, in the sense that when prices are falling the IP-IQ method tends also to miss price declines. … Prices have generally been falling for electronic products, including IT products. When the IP-IQ method is used to construct price indexes for electronic products, the price indexes are biased upward because they do not adequately measure price declines that accompany new introductions". The example here suggests that

[2] This is also often referred to as the "deletion" method.

what matters for the IP-IQ bias in IT product inflation is not whether prices are *declining over time* but rather whether prices per quality unit are *declining in the amount of quality embodied* in the models.

Hedonic price indexes are, in some sense, the opposite of IP-IQ matched model indexes. That is, where the IP-IQ method 'deletes' the observed prices of the unmatched models, hedonic methods 'insert' the unobserved prices of the unmatched models. This insertion, or more correctly 'imputation', is done by estimating a hedonic price equation that relates the price of a model in a particular period to a set of its quality characteristics and then using this equation to predict what the unobserved prices of the unmatched models would have been.

Over the past five years, hedonic price indexes have been implemented for an increasing number of goods for U.S. aggregate statistics. See Landefeld and Grimm (2000) as well as Moulton (2001), for example, for a discussion of the application of hedonic price indexes in the U.S. national accounts. The main reason why hedonic price indexes are adopted for an increasing number of goods is the practical problem that the IP-IQ method ends up not using a large part of the available price quotes in markets where new and obsolete models make up the bulk of models traded. This is particularly a problem for computers and related equipment.

The believe is that by taking the price data for the obsolete and new models into account and relating them directly to quality characteristics, hedonic price indexes more properly adjust for quality and are less subject to quality bias. This seems to be confirmed by the fact that hedonic price indexes tend to find less inflation for most of the goods to which they are applied[3] than standard matched model indexes, which are said to overestimate inflation.

Is it true that hedonic price indexes have a smaller quality bias than matched model indexes? Not necessarily. In order to see why not, consider Figure 3. Which prices are imputed in a hedonic price index depends on the price index formula applied. The two panels of Figure 3 depict the two most common cases.

The top panel considers a hedonic Laspeyres index, which intends to measure the percentage change in the cost of the models sold in period *0*. The Laspeyres index requires the use of the prices of the models that became obsolete in period *1*. Therefore, a hedonic regression model is used to impute these prices and the price schedule in period *1* is extended by the imputed part D_1-E_1. The Laspeyres index then basically approximates the change in the overall price levels implied by the curves A_0-D_0 and B_1-E_1. The overall price level implied by A_0-D_0 coincides with the actual price level in period *0*, i.e. P_0^T. The price level implied by B_1-D_1, denoted by P_1^{HL} in the figure, is lower than the actual price level in period *1*. The reason is that for the calculation of the Laspeyres index the above average prices per quality unit in the part A_1-B_1 are ignored. Moreover, the imputation adds below average prices per quality unit in the section D_1-E_1. Hence, the inferred price level in period *1* is below the actual price level and inflation is underestimated. In fact, because the A_0-D_0 schedule is above the B_1-E_1 schedule everywhere, in this example the hedonic price index would find spurious price deflation.

The bottom panel depicts the calculation of a hedonic Paasche index. It is meant to approximate the change in the cost of the models sold in period *1*. Therefore it requires the imputation of the D_0-E_0 part of the

[3] See for example Gordon's (1990) hedonic price indexes.

price schedule in period *0* and ignores the part A_0-B_0 in its calculation. The imputed part D_0-E_0 consists of above average prices per quality unit and the ignored part A_0-B_0 of below average prices per quality unit. This leads to the hedonic method overestimating the price level in period *0*. This estimate is denoted by P^{HP}_0. Again, the hedonic method will find spurious deflation. This time because it overestimates the price level in period *0*, rather than underestimates the price level in period *1*.

Thus, my tentative graphical example illustrates why matched model and hedonic methods might actually result in estimates of inflation that are *too low* rather than *too high*. However, this simple graphical example can only be used for illustration purposes, it does not prove that such biases might occur in the data. In order to show that these biases are likely to occur, I introduce a fairly standard theoretical model in the next section and show how the equilibrium outcome of the model gives rise to biases of the same kind as discussed here.

4. Theoretical model

The aim of this section is to introduce a simple theoretical framework that generates the kind of bias that I discussed in the section above. The theoretical framework introduced here is based on the CES model considered by Anderson, de Palma, and Thisse (1992). Feenstra (1995) applied this model to hedonic price indexes. I will introduce the theoretical model in three subsections. The first explains the demand side of the market, while the second focuses on the supply side of the market. In the final subsection, I will prove existence and uniqueness of the Pure Strategy Nash equilibrium that determines prices and quantities in the market and will derive some of the relevant comparative statics for this equilibrium.

Demand side of the market

Aggregate demand in this market can be represented as generated by a representative agent choosing the demand $\{X_{i,t}\}_{i \in C_t}$ to maximize the aggregate utility function in equation (1). This utility function is maximized subject to the budget constraint

$$Y_t = X_{0,t} + \sum_{i \in C_t^*} p_{i,t} X_{i,t} \tag{4}$$

where Y_t denotes real income in terms of the numeraire commodity and $p_{i,t}$ is the price of commodity i in terms of the numeraire good $X_{0,t}$.

The maximization of this utility function yields the demand functions

$$X_{i,t} = \left(\frac{\widetilde{Y}_t}{p_{i,t}} \right) \left(\frac{\left(p_{i,t} / a_i \right)^{-1/\lambda}}{\sum_{j \in C_t^*} \left(p_{j,t} / a_j \right)^{-1/\lambda}} \right) = \left(\frac{\widetilde{Y}_t}{p_{i,t}} \right) \left(\frac{\left(p_{i,t} / a_i \right)^{-1/\lambda}}{\left(P_t^T \right)^{-1/\lambda}} \right) \tag{5}$$

for the non-numeraire commodity, i.e. $i \in C^*_t$. The variable $\widetilde{Y}_t = Y_t / (1+\alpha)$ is the level of total expenditures on these models. These demand functions are very similar to the ones implied by standard CES preferences where the level of quality for all goods is the same, i.e. $a_i=1$ for all $i \in C^*_t$. The main difference is that the relevant relative price of each good that determines its market share is its price per unit of quality, that is p_{it}/a_i, rather than its unit price, p_{it}.

Supply side of the market

The next concern is the supply side of the market for $i \in C^*_t$. I will assume that the producer of model i at each point in time, t, faces a constant unit production cost c_{it}. I will consider Pure Strategy Nash equilibria in prices for a market with a fixed set of models, $\mathbf{a}_t = \{a_i\}_{i \in C_t^*}$. Such Nash equilibria imply that the supplier of model i takes as given the prices p_{jt} for $j \in C^*_t \backslash \{i\}$ and chooses its price p_{it} to maximize its profits

$$(p_{it} - c_{it}) X_{it} \tag{6}$$

subject to the demand function (5). The profit maximizing choice of price p_{it} in this case satisfies the following first order condition.

$$\frac{c_{it}}{p_{it}} = 1 + \left[\frac{\partial X_{it} / \partial p_{it}}{X_{it} / p_{it}} \right]^{-1} \tag{7}$$

This condition implies that the supplier of each model chooses its price such that its cost-price ratio equals one plus the inverse of the own price elasticity of demand for good i.

Since the own price elasticity of demand for good i is negative, this implies that $c_{it}/p_{it} < 1$. That is, price exceeds marginal and average cost and the firm charges a markup. For the price elasticity of demand, we obtain that

$$\frac{\partial X_{it} / \partial p_{it}}{X_{it} / p_{it}} = -\left(1 + \frac{1}{\lambda} \left(1 - \frac{(p_{it}/a_{it})^{-1/\lambda}}{\sum_{j \in C_t^*} (p_{jt}/a_{jt})^{-1/\lambda}} \right) \right) = -\theta_i(\mathbf{p}_t, \mathbf{a}_t) \tag{8}$$

where $\theta_i(\mathbf{p}_t, \mathbf{a}_t) > 0$ is the negative of the price elasticity of demand for good i and $\mathbf{p}_t = \{p_{it}\}_{i \in C_t^*}$ is the sequence of prices charged in the market. Essential for the results that are to follow is that this elasticity is specific to good i. This is contrary to the setup of monopolistic competition that is often used to model imperfect competition in models with price rigidities, like in Hornstein (1993). These models generally consider a symmetric equilibrium in which each monopolistic competitor is too small to affect the aggregate price level and its own price elasticity of demand.

Using the notation above, the supplier of good i will set its price such that

$$\frac{p_{it}}{c_{it}} = \frac{1}{1 - \theta_i(\mathbf{p}_t, \mathbf{a}_t)^{-1}} = \frac{\theta_i(\mathbf{p}_t, \mathbf{a}_t)}{\theta_i(\mathbf{p}_t, \mathbf{a}_t) - 1} = \mu_i(\mathbf{p}_t, \mathbf{a}_t) \tag{9}$$

where $\mu_i(\mathbf{p}_t, \mathbf{a}_t) > 1$ is the markup charged by the firm. Solving for this markup yields that

$$\mu_i(\mathbf{p}_t,\mathbf{a}_t) = (1+\lambda) + \lambda \frac{(p_{it}/a_i)^{-1/\lambda}}{\sum_{j\in C_t^*\setminus\{i\}} (p_{jt}/a_j)^{-1/\lambda}} \tag{10}$$

This implies that the pure strategy Nash equilibrium in this market satisfies the following system of equations

$$p_{it} = \left[(1+\lambda) + \lambda \frac{(p_{it}/a_i)^{-1/\lambda}}{\sum_{j\in C_t^*\setminus\{i\}} (p_{jt}/a_j)^{-1/\lambda}} \right] c_{it} \text{ for all } i \in C^*_t \tag{11}$$

This system of equations will be the center of attention in what is to follow.

Equilibrium

Now that I have derived the conditions for a Pure Strategy Nash equilibrium in equation (11), the question that remains is whether there exists a set of prices the satisfies this equation. In this section I will not only show that this is the case, but also prove the uniqueness of this equilibrium price schedule. I will then proceed by deriving some of its comparative statics that are relevant for the price index measurement results that I will prove later on.

First and foremost though, it is important to realize that the Pure Strategy Nash equilibrium in prices that I consider actually exists and is unique. This is what I prove in the following proposition.

Proposition 1: ***Existence and uniqueness of equilibrium***

For any $\lambda>0$ and sequences $\mathbf{a}_t = \{a_i\}_{i\in C_t^*}$ and $\mathbf{c}_t = \{c_{it}\}_{i\in C_t^*}$ where $a_i,c_{it}>0$ for all $i\in C^*_t$ there exists a unique Pure Strategy Nash equilibrium in prices.

The benchmark case, and as it turns out the only one in which standard price index methods do not generate a bias, is the case in which each supplier faces the same unit production cost per quality unit. As I show in the proposition below, the price per quality unit is the same for all models in the market in that case.

Proposition 2: ***Symmetric equilibrium***

The market has a symmetric equilibrium in which the price per quality unit is constant across models, i.e. $p_{it}=p^*_t a_i$ *for all* $i\in C^*_t$*, if and only if the producer of each model faces the same marginal unit production cost per quality unit, i.e.* $c_{it}=c^*_t a_i$.

In the previous section I argued that the bias that I illustrated graphically was the result of the price per quality unit not being constant across models sold in the market. In fact, in the example, the price per quality unit was higher for better models. In the symmetric equilibrium derived above the price per quality unit is constant and it is thus unlikely that this equilibrium will yield a bias of the sort described before. However, if the marginal production cost per quality unit is not the same across models sold in the market, then neither is the price charged per quality unit. In that case the market equilibrium will be asymmetric in the sense the models will have different market shares. As I show in the following proposition, the suppliers that produce the models with the higher marginal production cost per quality unit will charge a higher price per quality unit and will have a lower market share.

Proposition 3: ***Properties of asymmetric equilibrium***

In the asymmetric equilibrium, producers with higher marginal production costs per efficiency units, i.e. c_{it}/a_i, *(i) charge a higher price per efficiency unit,* p_{it}/a_i, *and (ii) a lower markup,* p_{it}/c_{it}.

The above result is important because it suggests that any asymmetric equilibrium exhibits prices per quality unit that are unequal across the models sold in the market and thus has the potential of generating the bias described in the previous section.

5. Price index bias in the theoretical model

Now that I have developed the theoretical model of this market, it is time to consider what conventional price index methods would measure in this market. In order to illustrate the quality bias it is essential to consider dynamic universes of goods such that

$$C_0^* \neq C_1^* \text{ and } C_0^* \cap C_1^* \neq \varnothing \tag{12}$$

In principle, there are many ways in which the set of models traded in the market can change and each of these changes might have a different effect in the theoretical example considered here. Because it is simply impossible to consider all of these different cases, I will limit myself to one specific example. In the first subsection, I will describe the parameterization of this example in detail. Then, in the second subsection, I will consider what happens when standard price index methods are applied in this example.

Parameterization of example

The example that I will consider is one where the model at 'the bottom of the line' in period $t=0$ becomes obsolete in period $t=1$ and in which in period $t=1$ a new 'top of the line' model is introduced. The 'bottom of the line' model at $t=0$ is the lowest-quality model, i.e. the one with the smallest a_i among all $i \in C^*_0$. The 'top of the line' model introduced in period $t=1$ is such that its quality exceeds that of all models traded in period $t=0$.

Consequently, in both periods the same number of models is sold[4]. I will denote this number by $N=N_0=N_1$. I will index the models as $i=1,\ldots,N+1$, where model 1 is the 'bottom of the line' model that becomes obsolete at time $t=1$ and model $N+1$ is the new 'top of the line' model introduced at $t=1$. This indexation implies that $C^*_0=\{1,\ldots,N\}$ and $C^*_1=\{2,\ldots,N+1\}$.

[4] Price indexes also have a problem measuring the increased utility from the availability of more models. This is known as variety bias. I will abstract from variety bias throughout this paper.

Two things are still to be defined. The first is the parameterization of the quality levels $\{a_i\}_{i=1}^{N+1}$. I will assume that quality is increasing in *i* such that

$$a_i = (1+g)^{i-1} \tag{13}$$

where *g>0* represents the quality growth rate across models.

The second is the parameterization of the unit production costs $\{c_{it}\}_{i=1}^{N+1}$. The parameterization that I will choose for these unit production costs is

$$c_{it} = c_t^* a_i^{1+\gamma} \tag{14}$$

This parameterization is such that if $\gamma=0$ then the production costs per quality unit are identical across models and the equilibrium is symmetric. If $\gamma<0$ then the production costs per quality unit are lower for better models and their suppliers will charge a lower price per quality unit and a higher markup in equilibrium, as shown in proposition 3. Similarly, if $\gamma>0$ then production costs per quality unit are higher for better models and, as in the graphical example, the price per quality unit is higher for better models. Hence, γ represents the steepness of the cross-model production costs per quality unit schedule. Because of proposition 3, this implies that γ also represents the steepness of the cross-sectional price per quality unit schedule.

I will parameterize the change of c^*_t over time as follows. Let $\overline{a}_t = \prod_{i \in C_t^*} a_t^{1/N}$ then I will assume that

$$c_t^* = \frac{\widetilde{c}_t}{\overline{a}_t^{\gamma}} \text{ where } \widetilde{c}_1 = (1+\pi)\widetilde{c}_0 \tag{15}$$

The reason that I parameterize c^*_t like this is because, in equilibrium, the structural parameter π has a specific interpretation. This is proven in the following proposition.

Proposition 4: ***Interpretation of structural parameter π***

In equilibrium, the structural parameter π *equals the theoretical inflation rate, i.e.* $\pi=\pi^T$.

Given this parameterization, the question is how estimated inflation on the basis of the various price index methods depends on the underlying structural parameters, *N*, *g*, π, γ, and λ and how it compares to the actual level of inflation, π^T. This question is addressed in the next subsection.

Quality bias

Just like in the graphical example of section 3, I will first address the bias induced by matched model indexes and then consider hedonic price indexes in this theoretical model.

For the matched model price indexes I will solely consider the, most frequently used, IP-IQ method. The following proposition states the properties of the IP-IQ linked matched model indexes in this example.

Proposition 5: ***Matched model index properties***

An IP-IQ linked matched model index yields an estimate of inflation, π^M, that has the following properties:

(i) $\pi^M = \pi^T$ *if* $\gamma=0$.

(ii) $\pi^M > \pi^T$ *if* $\gamma<0$.

(iii) $\pi^M < \pi^T$ *if* $\gamma>0$.

This result does not depend on which of the price index formulas (except the Feenstra (1994) index which is exact) is applied.

This proposition is the formal mathematical proof of the informal argument that I stated with respect to matched model price indexes for the graphical example in section 3. That is, the sign and magnitude of the quality bias in matched model price indexes *does not depend on the sign and magnitude of the overall inflation rate*. Instead, *it depends on the cross-sectional behavior of prices per quality unit for the models sold* in the market.

That the bias does not depend on the sign and magnitude of the overall inflation rate follows directly from the fact that the result in proposition 5 does not depend on the structural parameter π. The dependence of the bias on the steepness of the cross-sectional schedule of prices per quality unit across models is implied by the bias in the matched model indexes only depending on the parameter γ.

That is, if $\gamma>0$ then, according to proposition 3, the price per quality unit is increasing in the level of quality embodied in the model. This is the case depicted in the graphical example of section 3 and is the case that yields a *downward* bias in the measured inflation rate. If $\gamma<0$ then the reverse is true.

So, how do hedonic price methods behave in the theoretical model here? This question can only be answered conditional on the behavior of the imputed price levels. I do so in the next proposition.

Proposition 6: ***Hedonic price index properties***

A hedonic price index yields an estimate of inflation, π^H, that has the following properties:

(i) $\pi^H = \pi^T$ *if* $\gamma=0$.

(ii) $\pi^H > \pi^T$ *if* $\gamma<0$, *if the imputed prices satisfy the property of the equilibrium price schedule that prices per quality unit are decreasing in the quality embodied in the model.*

(iii) $\pi^H < \pi^T$ *if* $\gamma>0$, *if the imputed prices satisfy the property of the equilibrium price schedule that prices per quality unit are increasing in the quality embodied in the model.*

Just like in proposition 5, this result does not depend on which of the price index formulas (except the Feenstra (1994) index which is exact) is applied.

The proof of the proposition above gives some interesting insights. First of all, the hedonic price indexes only yield an unbiased estimate of inflation whenever the equilibrium is such that the price per quality unit is constant across the models traded in the market. However, if the price per quality unit is constant across models, then matched model indexes will do just fine. In fact, if the price per quality unit is constant across the models sold in the market, then one can simply measure overall inflation by considering the percentage price change of a single model. That is, when the price per quality unit is constant across the models sold in the market quality bias is not an issue. This itself is an important observation.

Bils and Klenow (2001) for example use microdata from the Consumer Expenditure Survey to estimate the quality bias in the CPI for several durable consumption goods. They do so by estimating a structural model of durable goods consumption. In order to quantify quality growth in this model, however, they assume that independent of each household's expenditures on a particular durable consumption good, the price paid per quality unit is constant for all households. Hence, no matter what model the households are buying, they are assumed to pay the same price per quality unit. This means that Bils and Klenow (2001) implicitly assume that the price per quality unit is constant across models. However, if this identifying assumption would be true in the data then the BLS would have had no problem quantifying quality growth in the first place.

If the price per quality unit is constant, then relative prices represent relative quality differences. In that case the coefficients in the hedonic regression model will represent the marginal quality coefficients of the quality indicators. Feenstra (1995) shows that when these coefficients represent these marginal values, hedonic price indexes will work properly. In fact, for certain classes of preferences Feenstra (1995) derives exact hedonic indexes. However, when he considers the existence of markups he also observes that when this is not the case then the estimated hedonic regression coefficients might over- or underestimate the quality difference between the models.

This is the case when $\gamma>0$ and $\gamma<0$. In those cases hedonic regression coefficients do not only reflect the marginal quality differences between the models but also reflect the slope of the price per quality unit schedule.

Log-linearized approximation of the bias

In order to consider how the size of the bias depends on the parameters, it turns out to be illustrative to consider the log-linear approximation of the bias around the symmetric equilibrium derived in Proposition 2.

Proposition 7: ***Log-linearization of the bias***

The log-linear approximation of the equilibrium around the symmetric equilibrium of Proposition 2 yields that both matched model as well as hedonic price indexes are subject to a bias equal to -θg, where

$$\theta = \gamma\left[1 - \frac{N}{1+(1+\lambda)N(N-1)}\right] \tag{16}$$

and does not depend on the actual inflation rate π.

What constitutes this bias? The reason for this bias is that the price index methods can not distinguish between a movement in the price per quality unit schedule over time due to an actual change in the overall price level and a move in the schedule because the introduction of a new model shifts the relative competitive advantages (in production and the market) and thus prices of the models sold in the market.

In order to see this, consider Figure 4. It disentangles the various effects on the schedule of the logarithm of the price per quality as a function of the logarithm of the quality. That is, it graphically represents the first difference of (59) over time and the various things that influence it.

Consider model *i* at time $t=0$. It has price p_{i0}, such that the logarithm of its price per efficiency unit is $ln(p_{i0}/a_i)$. At time $t=0$ it is at point *A* on the log price per quality unit schedule. For expositional purposes, I have drawn this graph for $\pi<0$ and $\gamma>0$. The drop in the overall price level $\pi<0$ shifts the log of the price per quality unit of model *i* down from point *A* to point *B*. However, something else happens at the same time as well. That is the introduction of the new model $N+1$ and the exit of model *1*.

Because of the introduction of the new, superior, model and the fact that production costs are increasing in the quality embodied in the model, the production costs of model *i* relative to those of its competitors will drop. This allows the supplier of model *i* to charge a lower price per quality unit than in period $t=0$. In fact, because of the setup of the model, model $i+1$ takes over model *i*'s position in the relative quality ladder in period $t=1$. Therefore, in period $t=1$ model $i+1$ will be sold at the same relative price per quality unit that model *i* was sold at in period $t=0$. This is depicted in Figure 4 by the horizontal shift from *B* to *C*. The slope of the log price per quality unit schedule, i.e. θ, and the length of the horizontal shift, i.e. *g*, then jointly determine how far below $ln(p_{i0}/a_i)+\pi$ the logarithm of model *i*'s price per quality unit in period $t=1$, i.e. $ln(p_{i1}/a_i)$, ends up.

Hence, in terms of this Figure 4, the problem of the price index methods is that they do not distinguish between the actual change in the overall price level, depicted by the shift from *A-B*, and the effect of the shift in the relative qualities of the models due to the introduction of a new model, depicted by the movement from *B-D*.

6. Data based simulation: CPU prices

So far, the point that I made is purely theoretical and I haven't addressed its empirical relevance. This section is intended to do so. In this section I illustrate how the effects described above would lead to, mostly downward, bias in measured inflation for CPU prices.

Experiment setup

The approach that I will take is as follows. I will use weekly price data for CPU prices to obtain an empirical set of prices, p_{it}. Furthermore, I use data on benchmark test performance of these CPUs as a measure of their quality, a_i. Using these two datasources[5], for each week I have a set of data on price per quality unit, p_{it}/a_i. The data do not contain any information on market shares of the different processors. I will simulate these data myself under the assumption that demand in the market is generated by the same CES preferences used in the theoretical model, i.e. equation (1). This requires the choice of the elasticity of substitution $\sigma=(1+\lambda)/\lambda>1$.

[5] The data are described in detail in Appendix B.

The result is that for various elasticities of substitution, I simulate a sequence of weekly cross-sectional samples of CPU prices and CPU market shares that are consistent with the preferences in (1). The nice thing of this simulation is that I know exactly the implied path of the theoretical price level. This then allows me to compare the price path estimated using conventional price index method with the path of the theoretical price level and to assess the sign and magnitude of the bias induced by these index methods.

Before considering the details of the results of this simulation, it turns out to be illuminating to first consider the behavior of prices per quality unit, i.e. p_{it}/a_i, for the CPUs in the sample. The sample covers about one and a half years, starting in 10/28/01 and ending 03/17/03. Figure 5 depicts the empirical equivalent of panel (b) of Figure 1. It plots the price per efficiency unit, i.e. the price per benchmark unit, as a function of the benchmark ratings. As can be seen from Figure 5, the price per efficiency unit schedule is increasing in the quality of the processor, both at the beginning and the end of the sample. These periods are no exception. In fact, the price per benchmark unit is increasing in the benchmark rating for all weeks in the sample.

The general cross-sectional pattern of CPU price declines is as follows. The prices of CPUs with the highest benchmark ratings tend to decline the fastest, while those on the bottom end of the range of CPUs sold barely change. The small price changes of CPUs at the bottom end of the spectrum might be partly induced by the way the data are collected. Probably, prices of cheap CPUs are still quoted for a while after they are sold. Anyway, because of this pattern of price declines we know beforehand that methods that weigh the price changes of the bottom end models more will find less price declines than methods that put a higher weight on the upper-end models.

The two observations above suggest the following. First, the fact that the price per benchmark unit is increasing in the benchmark rating creates the potential for price index methods to underestimate the price changes over the sample period. That is, quality bias for the CPU price index calculated based on these data is most likely *negative* rather than *positive*. Secondly, the different rates of price declines of low- and high-end CPUs suggest that price indexes that use different goods baskets might actually diverge over the sample period.

In order to consider how these effects manifest themselves in the data, I calculate the chained weekly CPU price indexes using the market shares implied by different values of the elasticity of substitution. The value for which the experiment is performed is 1.5, 2, 4, and 8. I have chosen to present two sets of results.

The first set calculates the price indexes for a varying number of models sold in the market. Because the CES preference of equation (1) satisfy Inada conditions, the availability of an additional CPU will always lower the price level. The effect of the number of goods on the price level is often referred to as the variety bias. The number of models in the weekly samples fluctuates between 37 and 50. The number of models at the beginning of the sample period was 46, which is 6 higher than at the end. Because of this, the number of models traded induces an upward trend in the price level that is not captured by the price indexes.

The second set of results calculates the price index for a constant number of models sold in the market. For these results I selected the 37 highest priced models in the market in each period and assumed that the market shares of the other models were zero. Because of this way of constructing the sample, this case does not suffer from the variety bias described above.

Results

Table 2 presents the results obtained for both cases. The first line of the table lists the simulated change in the actual price level over the sample period for the four different elasticities of substitution. As can be seen from this line, the decrease in the number of models reduces the magnitude of the implied price declines. That is, price declines with variety bias are smaller than those without variety bias.

The subsequent twelve lines list the estimated price changes over the sample period for both IP-IQ matched model indexes as well as hedonic price indexes using six commonly used price index formulas. These formulas are explained in detail in Table 1. I will first focus on the results for the matched model indexes and then discuss the results for the hedonic price indexes. Because the topic of this paper is the quality bias rather than variety bias in price indexes, I will mainly rely on the results without variety bias in the following.

When one considers the estimated price changes using the hedonic price indexes for the case without variety bias, one thing is immediately obvious. That is that for most elasticities of substitution all the price index formulas estimate a price decline that is larger than actually observed for the theoretical price level. That is, contrary to the consensus view, all the price indexes tend to underestimate CPU price deflation.

The downward bias is actually quite large. For the six price index formulas applied, the downward bias in the index varied between 18% and 24% in case of an elasticity of substitution of 1.5, between 12% and 21% for an elasticity of substitution of 2, between 1% and 20% for the elasticity of substitution of 4 and between –12% and 24% for the elasticity of substitution of 8. The downward bias in the price indexes is the sum of two different biases, namely the substitution bias and the quality bias.

The substitution bias has been extensively studied in the literature. Ever since Konüs (1939) and Frisch (1936) it is well known that price indexes based on the goods basket at the beginning of the period, like the Laspeyres and Geometric (G0) indexes, tend to overestimate inflation because they do not take into account the possibility of substituting away from the initial goods basket in response to relative price changes. On the other hand, price indexes that are based on the basket of goods at the end of the period, like the Paasche and Geometric (G1) indexes, tend to underestimate the actual inflation rate because they do not take into account the possibility of substituting away from the final goods basket in the initial period if relative prices made this desirable.

In order to minimize the potential substitution bias, superlative price indexes, like the Fisher and Tornqvist indexes, are often used. These indexes are exact indexes for a second order approximation to any arbitrary continuously differentiable utility function.

For the results in Table 2 it is important to realize that the substitution bias does not always work in the same direction. It pushes up the CPU price inflation estimated by the Laspeyres and Geometric (G0) indexes, because of which these find smaller price declines than the other index formulas. I pushes down the CPU price inflation estimate by the Paasche and Geometric (G1) indexes, because of which these tend to find larger CPU price declines than the other index formulas. For the superlative indexes, i.e. the Fisher and Tornqvist indexes, the substitution bias should be relatively small. Therefore, downward bias in the inflation rate estimated by these indexes can reasonably be solely attributed to the quality bias.

Taking the biases in the matched model Fisher and Tornqvist indexes as an estimate of the quality bias suggests that the quality bias in the matched model CPU price index is negative and large. It varies between approximately -20% for an elasticity of substitution of 1.5 to -9% for an elasticity of substitution of 8. This implies that a big chunk of the measured price declines in matched model CPU price indexes is spurious and due to quality bias.

It turns out that hedonic price indexes do not fare any better than the matched model indexes.

The hedonic price index results presented in Table 2 are calculated using imputed prices based on weekly log-log regressions in which the log of the CPU price is regressed on an intercept, an Athlon dummy, the logarithm of the clock speed (if applicable), the logarithm of the Athlon speed variable (if applicable), the logarithm of the bus speed, and the logarithm of the manufacturing process wiring width. The definition of these variables is explained in more detail in Appendix B. The estimated parameters for these hedonic regressions turn out to be fairly constant over the sample period and the R^2 varies between 0.69 and 0.83. The coefficients on the bus speed and the manufacturing process variables are insignificant for almost all of the sample period. Thus, the regression results seem to suggest that it is clock speed that is the driving force behind CPU prices.

This specification of the hedonic regression results in the imputation of prices that imply relatively high price declines for models that become obsolete in a period compared to other models on the lower end. Because of this, the hedonic Laspeyres and Geometric (G0) indexes, which weigh such obsolete models in their index calculation, find higher price declines than their matched model counterparts. The opposite is true for the imputed prices for new models at the high-end of the range of models sold. Consequently, the hedonic Paasche and Geometric (G1) indexes find less price declines than their matched model counterparts. On balance, though, the hedonic measures find slightly larger price declines than the matched model indexes. This can be seen from the two superlative indexes, i.e. the Fisher and Tornqvist.

Thus, just like in many other applications, for the example of CPU prices here the hedonic price indexes tend to measure less inflation than the matched model indexes. Although, the difference is not very large. However, because they measure less inflation, this does not imply that they measure inflation better. In fact, they are subject to a bigger quality bias than matched model indexes.

This is an important result because a large part of the literature on investment specific technological change[6] uses Gordon's (1990) hedonic equipment price index as a measure of the path of the 'true' quality adjusted equipment price. This is based on the premise that standard price indexes tend to overestimate inflation and that hedonic price indexes do a better job adjusting for quality. Consequently, the smaller equipment price inflation implied by Gordon's index relative to more standard indexes like those published by the Bureau of Economic Analysis is interpreted as quality improvements that go unmeasured in the standard indexes. The results in this paper suggest that the premise on which these studies are based might not be correct and that hedonic price indexes might actually exhibit a larger negative quality bias than matched model indexes.

---

[6] See, for example, Greenwood, Hercowitz, and Krusell (1997), Cummins and Violante (2001), and Violante, Ohanian, Ríos-Rull, and Krusell (2000).

7. Conclusion

In this paper I argued that the quality bias in price indexes does not necessarily always bias them upwards. I illustrated how the sign and the magnitude of this bias depend on the cross-sectional behavior of prices per quality unit across the models sold in the market. I did so by introducing a theoretical model that generated a quality bias in inflation as measured using the most common price index methods. The three main points that can be taken away from the analysis here are.

First and foremost, the quality bias can be both positive and negative. The sign of the bias does *not* depend on the actual underlying overall inflation rate. Instead, it solely depends on the cross-sectional behavior of prices per quality unit.

Secondly, the bias does not depend on which of the many proposed price index formulas are used to calculate the index. Laspeyres, Paasche, Geometric mean, Fisher Ideal, Tornqvist, and Sato indexes all performed is a similar manner in the theoretical model in this paper.

Finally, hedonic price indexes do not necessarily reduce the quality bias. In the examples in this paper, hedonic methods did just as poorly as matched model indexes. However, other examples, like the one given in Hobijn (2001), suggest that they might actually do worse in some cases.

This result is important because the application of hedonic price indexes seems to gain momentum both with statistical agencies, see Moulton (2001) for example, as well as with researchers. In fact, an extensive recent research agenda, including Greenwood, Hercowitz, and Krusell (1997), Violante, Ohanian, Rios-Rull, and Krusell (2000), and Cummins and Violante (2002), has been using Gordon's (1990) hedonic equipment price index as a measure of the 'true' quality adjusted price change for equipment in the U.S.. However, the results here suggest that one has to be careful in using this hedonic price index as such a benchmark. Simply because it measures less equipment price inflation than price indexes published by the Bureau of Economic Analysis and Bureau of Labor Statistics, does not necessarily mean it adjusts better for quality.

The results in this paper provide additional insights in which type of competitive circumstances are suspect to generating a bias, up or down, in the price indexes we calculate. Future research could focus on empirical tests of these conditions and on identifying in which markets what bias is the most likely to occur. At least it seems that the conventional wisdom that the quality bias biases measured inflation upward deserves a more thorough empirical verification.

References

Aizcorbe, Ana M., Carol Corrado, and Mark Doms (2000), "Constructing Price and Quantity Indexes for High Technology Goods", *mimeo*, Federal Reserve Board of Governors.

Aizcorbe, Ana M., and Patrick C. Jackman (1993), "The Commodity Substitution Effect in CPI Data", *Monthly Labor Review*, December 1993, 25-33.

Aizcorbe, Ana M., (2001), "Why Are Semiconductor Prices Falling So Fast? Industry Estimates and Implications for Productivity Measurement", *working paper*, Federal Reserve Board of Governors.

Anderson, Simon P., Andre de Palma, and Jacques-Francois Thisse (1992), *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press.

Bils, Mark, and Peter J. Klenow (2001), "Quantifying Quality Growth", *American Economic Review*, 91, 1006-1030.

Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Grilliches, Dale Jorgenson (1996), "Toward A More Accurate Measure of the Cost of Living, *Final Report of the Advisory Commission to Study The Consumer Price Index*.

Braithwait, Steven D. (1980), "The Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost of Living Indexes", *American Economic Review*, 70, 64-77.

Bureau of Labor Statistics (2002), "An Introductory Look at the Chained Consumer Price Index", *webpage:* http://www.bls.gov/cpi/ccpiintro.htm.

Cummins, Jason G., and Giovanni Violante (2002), "Investment Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences", *Review of Economic Dynamics*, 5, 243-284.

Diewert, Erwin (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics*, 4, 115-145.

Diewert, Erwin (2001), "Hedonic Regressions: A Consumer Theory Approach", *mimeo*, University of British Columbia.

Dixit, Avinash K., and Joseph E. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity", *American Economic Review*, 67, 297-308.

Feenstra, Robert C. (1994), "New Product Varieties and the Measurement of International Prices", *American Economic Review*, 84, 157-177.

Feenstra, Robert C. (1995), "Exact Hedonic Price Indexes", *Review of Economics and Statistics*, 77, 643-653.

Fisher, Irving (1922), *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Frisch, Ragnar (1936), "Annual Survey of General Economic Theory: The Problem of Index Numbers", *Econometrica*, 4, 1-38.

Gordon, Robert (1990), *The Measurement of Durable Goods Prices*, Chicago: University of Chicago Press.

Greenwood, Jeremy, Zvi Hercowitz and Per Krusell (1997), "Long-Run Implications of Investment-Specific Technological Change", *American Economic Review*, 87, 342-362.

Hobijn, Bart (2001), "Is Equipment Price Deflation A Statistical Artifact?", *Staff Report 139*, Federal Reserve Bank of New York.

Hornstein, Andreas (1993), "Monopolistic Competition, Increasing Returns to Scale, and The Importance of Productivity Shocks", *Journal of Monetary Economics*, 31, 299-316.

Hulten, Charles R. (2002), "Price Hedonics: A Critical Review", *mimeo*, University of Maryland.

Irwin, Douglas and Peter Klenow (1994), "Learning by Doing Spillovers in the Semiconductor Industry", *Journal of Political Economy*, 102, 1200-1227.

Konüs, A.A. (1939), "The Problem of the True Index of the Cost of Living", *Econometrica*, 7, 10-29.

Landefeld, J. Steven, and Bruce T. Grimm (2000), "A Note on the Impact of Hedonics and Computers on Real GDP", *Survey of Current Business*, December 2000, 17-22.

Lebow, David E. and Jeremy B. Rudd (2003), "Measurement Error in the Consumer Price Index: Where Do We Stand?", *Journal of Economic Literature*, 41, 159-201.

Lloyd, P.J. (1975), "Substitution Effects and Biases in Nontrue Price Indices", *American Economic Review*, 65, 301-313.

Manser, Marilyn E. and Richard J. McDonald (1988), "An Analysis of Sustitution Bias in Measuring Inflation, 1959-85", *Econometrica*, 56, 909-938.

Moulton, Brent R. (2001), "The Expanding Role of Hedonic Methods in the Official Statistics of the United States", *working paper*, Bureau of Economic Analysis.

Pakes, Ariel (2002), "A Reconsideration of Hedonic Price Indices with an Application to PCs", *NBER working paper 8715*, National Bureau of Economic Research.

Sato, Kazuo (1976), "The Ideal Log-Change Index Number", *The Review of Economics and Statistics*, 58, 223-228.

Silver, Mick and Saeed Heravi (2002), "The Measurement of Quality Adjusted Price Changes", *working paper*, Cardiff Business School.

Triplett, Jack E. (1972), "Quality Bias in Price Indexes and New Methods of Quality Measurement", in *Price Indexes and Quality Change: Studies in New Methods of Measurement*, Zvi Grilliches (ed.), Cambridge, MA: Harvard University Press.

Triplett, Jack E. (2002), "Quality Adjustments in Conventional Price Index Methodologies", *Handbook on Quality Adjustment of Price Indexes For Information and Communication Technology Products*, OECD: Paris.

Violante, Giovanni, Lee Ohanian, José-Victor Ríos-Rull, and Per Krusell (2000), "Capital-skill Complementarities and Inequality: A Macroeconomic Analysis", *Econometrica*, 68, 1029-1053.

A. Proofs of propositions

Proof of proposition 1: (*existence*) Existence of the pure strategy Nash equilibrium follows from the application of Brouwer's fixed point theorem. In order to see how Brouwer's fixed point theorem applies here, it is most convenient to define $c^*_{it}=c_{it}/a_i$. Rewrite the system of equations that defines the equilibrium, i.e. (11), in the form

$$\left(\frac{p_{it}}{c^*_{it}a_i}\right)=\left[1+\lambda\frac{\sum_{j\in C^*_t}\left(c^*_{jt}\right)^{-1/\lambda}\left(p_{jt}/c^*_{jt}a_j\right)^{-1/\lambda}}{\sum_{j\in C^*_t}\left(c^*_{jt}\right)^{-1/\lambda}\left(p_{jt}/c^*_{jt}a_j\right)^{-1/\lambda}-\left(c^*_{it}\right)^{-1/\lambda}\left(p_{it}/c^*_{it}a_i\right)^{-1/\lambda}}\right] \text{ for all } i\in C^*_t \tag{17}$$

which implies that

$$\left(\frac{p_{it}}{c_{it}}\right)^{-1/\lambda}=\left[1+\lambda\frac{\sum_{j\in C^*_t}\left(c^*_{jt}\right)^{-1/\lambda}\left(p_{jt}/c_{jt}\right)^{-1/\lambda}}{\sum_{j\in C^*_t}\left(c^*_{jt}\right)^{-1/\lambda}\left(p_{jt}/c_{jt}\right)^{-1/\lambda}-\left(c^*_{it}\right)^{-1/\lambda}\left(p_{it}/c_{it}\right)^{-1/\lambda}}\right]^{-1/\lambda} \text{ for all } i\in C^*_t \tag{18}$$

Let N_t be the number of elements of C^*_t, i.e. the number of competing models in the market at time t. Define

$$v_{it}=\left(p_{it}/c_{it}\right)^{-1/\lambda} \text{ and the space } V_t=\left\{\{v_{it}\}_{i\in C^*_t}\in\mathbf{R}^{N_t}\,\middle|\,0\le v_{it}\le 1\right\} \tag{19}$$

then (18) defines a continuous mapping from V_t to V_t and thus, according to Brouwer's fixed point theorem must have a fixed point. Hence, there must exist an equilibrium.
(uniqueness) Define

$$z_{it}=\left(c^*_{it}\right)^{-1/\lambda},\ w_{it}=z_{it}\left(\frac{p_{it}}{c_{it}}\right)^{-1/\lambda}, \text{ and } W_t=\sum_{i\in C^*_t}w_{it} \tag{20}$$

then (18) can we rewritten as

$$w_{it}=z_{it}\left[1+\lambda\frac{W_t}{W_t-w_{it}}\right]^{-1/\lambda} \text{ for all } i\in C^*_t \tag{21}$$

Given W_t, for all $i\in C^*_t$, there is one unique $w_{it}\in\mathbf{R}_+$ that solves (21). This follows from a straightforward application of the intermediate value theorem to (21). Define the function $f\colon [0,W_t]\to\mathbf{R}_+$ as

$$f\left(w_{it}\right)=w_{it}-z_{it}\left[1+\lambda\frac{W_t}{W_t-w_{it}}\right]^{-1/\lambda} \tag{22}$$

then $f(w_{it})$ is continuous and strictly increasing. Furthermore, $f(0)=-z_{it}[1+\lambda]^{-1/\lambda}$ and $f(W_t)=W_t$. Hence, the intermediate value theorem implies that there must be a unique $w_{it,}\in[0,W_t]$ for which $f(w_{it})=0$.

Suppose the equilibrium is not unique, then there exist W_t and W'_t such that $W_t > W'_t = (1+\delta)W_t$, where $\delta > 0$, such that

$$W_t = \sum_{i \in C_t^*} w_{it} \text{ and } W'_t = \sum_{i \in C_t^*} w'_{it} \tag{23}$$

and W_t and w_{it} for all $i \in C^*_t$ satisfy (21), which is also true for W'_t and w'_{it} for all $i \in C^*_t$.

Note that the reason that W_t and W'_t can not be the same is because I have shown above that the same W_t will lead to the same best response by the suppliers of all models and thus to the same equilibrium.

What I will show in the following is that if (21) holds for W_t and w_{it} for all $i \in C^*_t$, then for all $i \in C^*_t$ it must be the case that the w'_{it} that satisfies (21) given W'_t has to satisfy $w'_{it} < (1+\delta)w_{it}$. However, this would imply that $W'_t < (1+\delta)W_t = W'_t$ which is a contradiction.

In order to see this, suppose that $w'_{it} \geq (1+\delta)w_{it}$. In that case, equation (21) implies that

$$\begin{aligned} w'_{it} &= z_{it}\left[1+\lambda\frac{W'_t}{W'_t - w'_{it}}\right]^{-1/\lambda} = z_{it}\left[1+\lambda\frac{(1+\delta)W_t}{(1+\delta)W_t - w'_{it}}\right]^{-1/\lambda} \\ &\leq z_{it}\left[1+\lambda\frac{(1+\delta)W_t}{(1+\delta)W_t - (1+\delta)w_{it}}\right]^{-1/\lambda} = z_{it}\left[1+\lambda\frac{W_t}{W_t - w_{it}}\right]^{-1/\lambda} = w_{it} \\ &< (1+\delta)w_{it} \leq w'_{it} \end{aligned} \tag{24}$$

which is a contradiction. Hence, there can only be one equilibrium. ■

Proof of proposition 2: (⇒) If $c_{it}/a_i = c^*_t$ and does not depend on i, then (11) reduces to

$$\left(\frac{p_{it}}{a_i}\right) = \left[1+\lambda+\lambda\frac{(p_{it}/a_i)^{-1/\lambda}}{\sum_{j \in C_t^* \setminus \{i\}}(p_{jt}/a_j)^{-1/\lambda}}\right]c_t^* \text{ for all } i \in C^*_t \tag{25}$$

When we choose $p_{it}/a_t = p^*_t$ for all $i \in C^*_t$ and substitute it in the system of equations (25) we obtain that for all $i \in C^*_t$

$$\begin{aligned} p_t^* &= \left[1+\lambda+\lambda\frac{p_t^{*-1/\lambda}}{\sum_{j \in C_t^* \setminus \{i\}} p_t^{*-1/\lambda}}\right]c_t^* \\ &= \left[1+\lambda\frac{N_t}{N_t - 1}\right]c_t^* \end{aligned} \tag{26}$$

which does not depend on i. Hence, if $c_{it}/a_i = c^*_t$, then $p^*_t = (1+\lambda N_t/(N_t-1))c^*_t$ is the symmetric pure strategy Nash equilibrium in which all suppliers charge the same price per quality unit and all have an equal market share.

(⇐) If there is a symmetric equilibrium, then for all $i \in C^*_t$

$$\left(\frac{p_{it}}{a_i}\right) = p_t^* = \left[1+\lambda\frac{N_t}{N_t-1}\right]\left(\frac{c_{it}}{a_i}\right) \tag{27}$$

which implies that

$$\left(\frac{c_{it}}{a_i}\right) = p_t^* \Big/ \left[1+\lambda\frac{N_t}{N_t-1}\right] = c_t^* \tag{28}$$

and does not depend on $i \in C^*_t$. ■

Proof of proposition 3: (*i*) Equation (25) implies that, when we define

$$W_t = \sum_{j\in C_t^*}\left(p_{it}/a_i\right)^{-1/\lambda}, \tag{29}$$

then for all $i \in C^*_t$ it must be in equilibrium that

$$\left(\frac{p_{it}}{a_i}\right) = \left[1+\lambda\frac{W_t}{W_t-\left(p_{it}/a_i\right)^{-1/\lambda}}\right]\left(\frac{c_{it}}{a_i}\right) \tag{30}$$

Applying the implicit function theorem to the above equation yields

$$\frac{\partial\left(p_{it}/a_i\right)}{\partial\left(c_{it}/a_i\right)} > 0 \tag{31}$$

Furthermore, because (p_{it}/a_i) is higher, equation (5) implies that the market share of the model must be lower. (*ii*) In order to prove this part, it is easiest to reconsider (18), which reads

$$\left(\frac{p_{it}}{c_{it}}\right)^{-1/\lambda} = \left[1+\lambda\frac{\sum_{j\in C_t^*}\left(c_{jt}^*\right)^{-1/\lambda}\left(p_{jt}/c_{jt}\right)^{-1/\lambda}}{\sum_{j\in C_t^*}\left(c_{jt}^*\right)^{-1/\lambda}\left(p_{jt}/c_{jt}\right)^{-1/\lambda}-\left(c_{it}^*\right)^{-1/\lambda}\left(p_{it}/c_{it}\right)^{-1/\lambda}}\right]^{-1/\lambda} \tag{32}$$

Again, redefining

$$v_{it} = \left(p_{it}/c_{it}\right)^{-1/\lambda} \text{ and defining } \widetilde{V}_t = \sum_{i\in C_t^*}\left(c_{it}^*\right)^{-1/\lambda} v_{it} \tag{33}$$

equation (32) boils down to

$$v_{it} = \left[1+\lambda\frac{\widetilde{V}_t}{\widetilde{V}_t-\left(c_{it}^*\right)^{-1/\lambda}v_{it}}\right]^{-1/\lambda} \quad \text{for all } i\in C^*_t \tag{34}$$

It is straightforward to show that the v_{it} that solves this equation is increasing in c^*_{it}. Since the markup, p_{it}/c_{it}, is decreasing in v_{it}, this implies that the equilibrium markup is decreasing in c^*_{it}, which is what is claimed. ■

Proof of proposition 4: The basis of this proof is the equilibrium equation

$$p_{it} = \left[(1+\lambda) + \lambda \frac{(p_{it}/a_i)^{-1/\lambda}}{\sum_{j\in C_t^*\setminus\{i\}} (p_{jt}/a_j)^{-1/\lambda}} \right] c_{it} \text{ for all } i\in C^*_t \tag{35}$$

when we define

$$\tilde{p}_{it} = \frac{p_{it}}{a_i \tilde{c}_t} \tag{36}$$

then (35) can be rewritten as

$$\tilde{p}_{it} = \left[(1+\lambda) + \lambda \frac{(\tilde{p}_{it})^{-1/\lambda}}{\sum_{j\in C_t^*\setminus\{i\}} (\tilde{p}_{jt})^{-1/\lambda}} \right] \left(\frac{a_i}{\overline{a}_t} \right)^{\gamma} \text{ for all } i\in C^*_t \tag{37}$$

what I will show is that

$$a_i / \overline{a}_t = a_{i+1} / \overline{a}_{t+1} \tag{38}$$

If this is the case, then it must be that $\tilde{p}_{it} = \tilde{p}_{(i+1)(t+1)}$. We know that

$$C_0^* = \{1,\ldots,N\} \text{ and } C_1^* = \{2,\ldots,N+1\} \tag{39}$$

Given these sets, (13) implies that

$$\overline{a}_0 = (1+g)^{\frac{1}{N}\sum_{i=1}^{N}(i-1)}, \text{ while } \overline{a}_1 = (1+g)^{\frac{1}{N}\sum_{i=2}^{N+1}(i-1)} = (1+g)\left[(1+g)^{\frac{1}{N}\sum_{i=1}^{N}(i-1)} \right] = (1+g)\overline{a}_0 \tag{40}$$

Since $a_{i+1}=(1+g)a_i$, this implies that

$$a_{i+1} / \overline{a}_{t+1} = (1+g)a_i / (1+g)\overline{a}_t = a_i / \overline{a}_t \tag{41}$$

Therefore, $\tilde{p}_{it} = \tilde{p}_{(i+1)(t+1)}$. This implies that

$$P_0^T = \left[\sum_{i=1}^{N} (p_{i0}/a_i)^{-1/\lambda} \right]^{-\lambda} = \tilde{c}_0 \left[\sum_{i=1}^{N} \tilde{p}_{i0}^{-1/\lambda} \right]^{-\lambda} \tag{42}$$

and

$$P_1^T = \left[\sum_{i=2}^{N+1} (p_{i1}/a_i)^{-1/\lambda} \right]^{-\lambda} = \tilde{c}_1 \left[\sum_{i=2}^{N+1} \tilde{p}_{i1}^{-1/\lambda} \right]^{-\lambda} = \tilde{c}_1 \left[\sum_{i=1}^{N} \tilde{p}_{(i+1)1}^{-1/\lambda} \right]^{-\lambda} = \frac{\tilde{c}_1}{\tilde{c}_0} P_0^T \tag{43}$$

Hence,

$$\pi^T = \frac{P_1^T - P_0^T}{P_0^T} = \frac{\widetilde{c}_1 - \widetilde{c}_0}{\widetilde{c}_0} = \pi \tag{44}$$

Thus, π represents the theoretical inflation rate that is supposed to be approximated by the empirical price index methods. ■

Proof of proposition 5: I will prove part *(iii)* of the proposition in detail. The other two parts follow directly from the proof below. Note that the inflation rate of good *i* between *t=0* and *t=1* is given by

$$\begin{aligned}\pi_i &= \frac{p_{i1} - p_{i0}}{p_{i0}} = \frac{\widetilde{c}_1 \widetilde{p}_{i1} - \widetilde{c}_0 \widetilde{p}_{i0}}{\widetilde{c}_0 \widetilde{p}_{i0}} \\ &= \frac{\widetilde{c}_1 - \widetilde{c}_0}{\widetilde{c}_0} + \frac{\widetilde{c}_1}{\widetilde{c}_0}\left[\widetilde{p}_{i1} - \widetilde{p}_{i0}\right] \\ &= \pi^T + \left(1+\pi^T\right)\left(\widetilde{p}_{i1} - \widetilde{p}_{i0}\right)\end{aligned} \tag{45}$$

Hence, what will be essential in the rest of this proof is the property of $\widetilde{p}_{i1} - \widetilde{p}_{i0}$. It turns out that $\widetilde{p}_{it}$ is increasing in $\left(a_i / \overline{a}_t\right)^\gamma$. In order to see why, it is useful to rewrite (37) as

$$\widetilde{p}_{it} = \left[1 + \lambda \frac{\widetilde{P}_t}{\widetilde{P}_t - \left(\widetilde{p}_{it}\right)^{-1/\lambda}}\right]\left(\frac{a_i}{\overline{a}_t}\right)^\gamma \quad \text{for all } i \in C^*_t \tag{46}$$

where

$$\widetilde{P}_t = \sum_{i \in C_t^*} \widetilde{p}_{it}^{-1/\lambda} \tag{47}$$

Applying the implicit function theorem to the above two equations yields in a straightforward manner that

$$\frac{\partial \widetilde{p}_{it}}{\partial \left(a_i / \overline{a}_t\right)^\gamma} > 0 \tag{48}$$

Therefore $\widetilde{p}_{it}$ is strictly increasing in $\left(a_i / \overline{a}_t\right)^\gamma$. Furthermore, note that if $\gamma=0$, then the equilibrium is symmetric and $\widetilde{p}_{it}$ is equal for all $i \in C^*_t$.

Consequently, if $\gamma>0$ then $\left(a_i / \overline{a}_t\right)^\gamma$ is increasing in *i* and models of higher quality have a higher $\widetilde{p}_{it}$. This implies that if $\gamma>0$ then

$$\widetilde{p}_{(i+1)0} > \widetilde{p}_{i0} = \widetilde{p}_{(i+1)1} \tag{49}$$

where the second equality follows from the proof of proposition 4. Hence, if $\gamma>0$ then for all $i \in C^*_t$,

$$\pi_i = \pi^T + \left(1+\pi^T\right)\left(\widetilde{p}_{i1} - \widetilde{p}_{i0}\right) < \pi^T \tag{50}$$

Because this inequality holds for all models sold in the market in periods *t=0* and *t=1*, matched model price indexes calculated using the Laspeyres, Paasche, Geometric mean, Fischer, Tornqvist, and Sato formula will all underestimate inflation. The reason is that all these price index formula have the property that measured inflation is in the range of inflation rates of the individual models. Since the actual inflation rate is above the maximum inflation rate measured for the models it must be that the actual inflation rate is understated by the matched model indexes.

A reverse but similar argument yields that the matched model indexes overstate inflation whenever $\gamma<0$. ■

Proof of proposition 6: *(i):* If $\gamma=0$ then the equilibrium price schedule satisfies

$$\frac{p_{it}}{a_i} = p_t^* \text{ for all } i\in C^*_t \text{ and for } t=0,1 \tag{51}$$

If the imputed prices, $\hat{p}_{N+1,0}$ and $\hat{p}_{1,1}$,from the hedonic regression model also satisfy this property such that

$$\frac{\hat{p}_{N+1,0}}{a_{N+1}} = p_0^* \text{ as well as } \frac{\hat{p}_{1,1}}{a_1} = p_1^* \tag{52}$$

then we find that the observed and imputed inflation rates satisfy

$$\pi_i = \begin{cases} \dfrac{p_{i,1}-p_{i,0}}{p_{i,0}} = \dfrac{p_1^*-p_0^*}{p_0^*} = \pi^T & \text{for } i = 2,\ldots,N \\ \dfrac{p_{N+1,1}-\hat{p}_{N+1,0}}{\hat{p}_{N+1,0}} = \dfrac{p_1^*-p_0^*}{p_0^*} = \pi^T & \text{for } i = N+1 \\ \dfrac{\hat{p}_{1,1}-p_{1,0}}{p_{1,0}} = \dfrac{p_1^*-p_0^*}{p_0^*} = \pi^T & \text{for } i = 1 \end{cases} \tag{53}$$

Consequently, no matter what type of weighted average one takes of the observed and imputed inflation rates across models to calculate π^H, this average will always equal π^T.

(ii): If $\gamma<0$ then the equilibrium price schedule satisfies

$$\frac{p_{it}}{a_i} < \frac{p_{i-1t}}{a_i} \text{ for all } i,i\text{-}1\in C^*_t \text{ and for } t=0,1 \tag{54}$$

If the imputed prices in the hedonic regression model also satisfy this property, such that

$$\frac{\hat{p}_{N+1,0}}{a_{N+1}} < \frac{p_{N,0}}{a_N} \text{ and } \frac{\hat{p}_{1,1}}{a_1} > \frac{p_{2,1}}{a_1} \tag{55}$$

then, in terms of the notation of proposition 5, the observed and imputed prices obey

$$\tilde{p}_{i,0} < \tilde{p}_{i-1,0} = \tilde{p}_{i,1} \text{ for } i=2,\ldots,N, \text{ as well as } \hat{\tilde{p}}_{N+1,0} < \tilde{p}_{N,0} = \tilde{p}_{N+1,1} \text{ and } \hat{\tilde{p}}_{1,1} > \tilde{p}_{2,1} = \tilde{p}_{1,0} \tag{56}$$

This means that the observed and imputed inflation rates satisfy

$$\pi_i = \begin{cases} \pi^T + (1+\pi^T)(\tilde{p}_{i,1} - \tilde{p}_{i,0}) > \pi^T & \text{for } i = 2,\ldots,N \\ \pi^T + (1+\pi^T)(\tilde{p}_{N+1,1} - \hat{\tilde{p}}_{N+1,0}) > \pi^T & \text{for } i = N+1 \\ \pi^T + (1+\pi^T)(\hat{\tilde{p}}_{1,1} - \tilde{p}_{1,0}) > \pi^T & \text{for } i = 1 \end{cases} \tag{57}$$

Hence, no matter what weighted average one takes of these inflation rates across models to calculate the hedonic inflation rate π^H, it will always yield $\pi^H > \pi^T$.

(iii): This follows in the same way as the proof of part *(ii)*. The only thing that is different is that in this case the equilibrium price schedule is such that the prices per quality unit are increasing in the quality levels of the models, which yields a reversal of the inequality signs. ■

Proof of proposition 7: Log-linearization of equation (30) around the symmetric equilibrium derived in Proposition 2 yields

$$\begin{aligned} \ln p_{it} &\approx \ln \tilde{c}_t + \ln\left(1 + \lambda \frac{N}{N-1}\right) + \ln a_i + \gamma\left[1 - \frac{N}{1+(1+\lambda)N(N-1)}\right](\ln a_i - \ln \bar{a}_t) \\ &= \left[\ln \tilde{c}_t + \ln\left(1 + \lambda \frac{N}{N-1}\right) - \gamma\left[1 - \frac{N}{1+(1+\lambda)N(N-1)}\right]\ln \bar{a}_t\right] + \left[1 + \gamma - \gamma \frac{N}{1+(1+\lambda)N(N-1)}\right]\ln a_i \\ &= \beta_{0t} + \beta_{1t} \ln a_i \end{aligned} \tag{58}$$

This implies that the logarithm of the price per quality unit satisfies

$$\ln(p_{it}/a_i) \approx \ln \mu_{it} + \ln \tilde{c}_t + \gamma(\ln a_i - \ln \bar{a}_t) \tag{59}$$

where $ln\mu_{it}$ is the logarithm of the markup charged on model i in period t which equals approximately

$$\ln \mu_{it} \approx \ln\left(1 + \lambda \frac{N}{N-1}\right) - \gamma\left[\frac{N}{1+(1+\lambda)N(N-1)}\right](\ln a_i - \ln \bar{a}_t) \tag{60}$$

When we take the first difference, over time, of (59), then we obtain that the percentage price change in the price of model i between $t=0$ and $t=1$ can be approximated by

$$\begin{aligned} \pi_i &\approx \Delta \ln p_{i1} \approx \Delta \ln \tilde{c}_1 + \Delta \ln \mu_{i1} - \gamma \Delta \ln \bar{a}_1 \approx \pi^T + \Delta \ln \mu_{i1} - \gamma \Delta \ln \bar{a}_1 \\ &\approx \pi^T - \gamma\left[1 - \frac{N}{1+(1+\lambda)N(N-1)}\right] g = \pi^T - \theta g \end{aligned} \tag{61}$$

Thus, we obtain that the inflation rates of each of the matched models deviate from the actual inflation rate by approximately the same amount, namely $-\theta g$. Because the hedonic regression extrapolates this approximately linear relationship for the imputation of the unobserved prices, it also finds that the imputed inflation levels $\pi_{N+1} = \pi_1 = \pi^T - \theta g$. Therefore, the results for the hedonic price indexes do not differ much from the matched model indexes and both are biased by approximately $-\theta g$. ■

B. Data on CPU prices and benchmark tests

The data are constructed from two main sources. The first contains the data on the CPU prices. The second contains the data on the benchmark test results and technical specifications of the various chips.

CPU price data: Weekly data on CPU prices for the period of October 28 2001 through March 17 2003 are taken from www.sharkyextreme.com. These prices are the lowest online list prices for CPU units of different types[7]. For comparability purposes, I have limited the sample to Original Equipment Manufacturer (OEM) processors, sometimes referred to as tray processors. These are processors that are sold to an OEM manufacturer or distributor intended for installation. These processors generally do not include a heat sink or fan and have a shorter warranty than their retail versions. The OEM versions tend to sell for about 15% less than the retail versions of the same processors.

These price data contain prices for 77 different processors. However, not all of these processors are sold in each week. The number of processors for which there are price quotes in a week varies between 37 and 50.

Benchmark test data: Quality of the CPU's is proxied for by benchmark ratings. The benchmark ratings that I use are taken from www.tomshardware.com[8]. They are taken from a comparative benchmark test of 65 processors varying in clock speed from 100MHz through 3066 MHz. I used data on 58 processors for which I was able to find detailed data on their technical specifications and which were classified as OEM versions. The processor in this sample with the lowest benchmark rating is the Intel Celeron 400MHz with a rating of 39. The highest rated processor, with a rating of 206, is the Pentium 4 3.06 MHz.

Technical specifications: Data on the technical specifications of the chips in my sample are taken from Tom's Hardware for the chips reported on in the benchmark test and from the Intel and AMD web sites for the other chips.

The technical specifications variables that I used in the benchmark imputation as well as the hedonic price regressions can be described as follows. *Athlon dummy*: Equals one if the processor is an AMD Athlon processor for which clock speed is not reported in MHz but instead in the measure AMD uses for speed, zero otherwise. *Speed:* Clock speed in MHz. *Athlon speed:* Clock speed measure published by AMD for most of its Athlon processors. *Bus speed:* Speed in MHz with which the processor communicates with the systems RAM memory. *Process:* Wiring width measured in nanometers used for the manufacturing process. The lower this width the higher the density of transistors on the chip can be.

Merger of price and benchmark data: The price data cover a set of processors that are not included in the benchmark test (and the benchmark test contains some processors that are not in the price data). I imputed a

---

[7] A detailed description of how SharkExtreme collects these data on CPU prices is available at http://www.sharkyextreme.com/guides/WCPG/article.php/10705_2224371__6.

[8] In particular, the benchmark tests that use are reported at http://www6.tomshardware.com/cpu/20030217/index.html.

benchmark rating for these untested processes using the following regression that aims to explain the benchmark rating as a function of processor specific technical characteristics.

$$\begin{aligned}\ln(\text{benchmark}) = \underset{(0.58)}{2.15} + \underset{(0.65)}{0.75}(\text{Athlon dummy}) \\ + \underset{(0.04)}{0.53}\ln(\text{Speed}) + \underset{(0.08)}{0.43}\ln(\text{Athlon speed}) \\ + \underset{(0.03)}{0.12}\ln(\text{Bus speed}) - \underset{(0.08)}{0.38}\ln(\text{Process})\end{aligned} \quad (62)$$

The sample of this regression consists of the 58 processors from the benchmark data and the R^2 equals 0.98. All technical specification variables come in significant and with the expected sign.

Combining the observed and imputed benchmarks with the price data yields a weekly data set with prices for 77 different processors for the approximately one and a half year between October 28 2001 and March 17 2003. The data set does not only contain data on prices but also on benchmark ratings and technical specifications of the various processors.

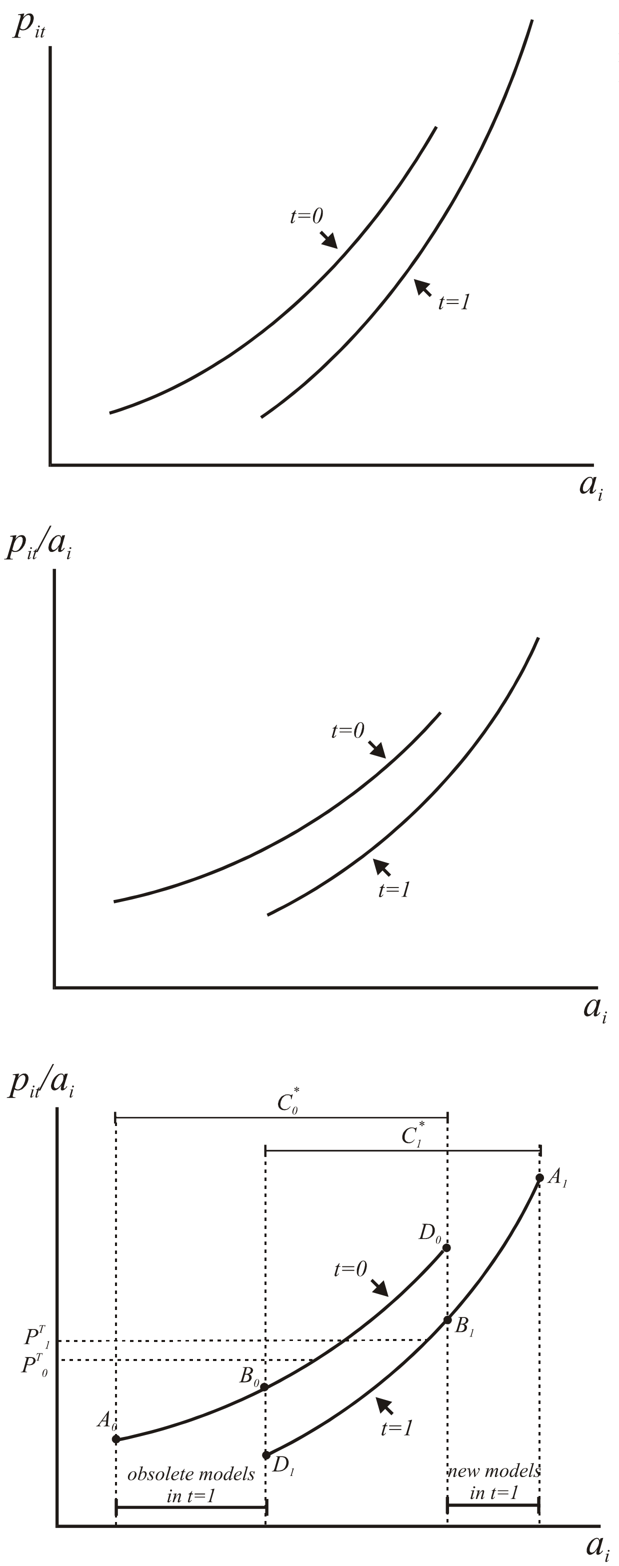
Table 1. Price index formulas applied in this paper

Index	Formula	Applied
Laspeyres	$\pi^L = \sum_{i \in C^*} s_{i0} \pi_i$	• Applied by Bureau of Labor Statistics for the calculation of the Consumer Price Index.
Paasche	$\pi^P = \left[\sum_{i \in C^*} s_{i1} (1+\pi_i)^{-1} \right]^{-1} - 1$	
Geometric mean (G0)	$\pi^{G_0} = \left(\prod_{i \in C^*} (1+\pi_i)^{s_{io}} \right) - 1$	
Geometric mean (G1)	$\pi^{G_1} = \left(\prod_{i \in C^*} (1+\pi_i)^{s_{i1}} \right) - 1$	
Fisher Ideal	$\pi^{FI} = \sqrt{(1+\pi^L)(1+\pi^P)} - 1$	• Applied by Bureau of Economic Analysis for the calculation of chained price indexes in the National Income and Product Accounts.
Tornqvist	$\pi^{TQ} = \sqrt{(1+\pi^{G_0})(1+\pi^{G_1})} - 1$	• Applied by Bureau of Labor Statistics (2002) in chained Consumer Price Index.
Sato	$\pi^S = \left(\prod_{i \in C^*} (1+\pi_i)^{\phi_i} \right) - 1$ where $\phi_i = \theta_i / \sum_{j \in C^*} \theta_j$ $\theta_i = \begin{cases} \frac{s_{i1} - s_{i0}}{\ln s_{i1} - \ln s_{i0}} & \text{if} \quad s_{i0} \neq s_{i1} \\ s_{i1} & \text{if} \quad s_{i0} = s_{i1} \end{cases}$	• Exact price index for CES preferences when same goods are sold in both periods.
Feenstra	$\pi^F = \left(\prod_{i \in (C_0^* \cap C_1^*)} (1+\pi_i)^{\phi_i} \right) \left(\frac{\varpi_1}{\varpi_0} \right)^{1/(1-\sigma)} - 1$ where ϕ_i is as in the Sato index $\varpi_t = \sum_{i \in (C_0^* \cap C_1^*)} s_{it}$ for $t=0,1$ σ elasticity of substitution ($\sigma=1+\lambda/\lambda$)	• Exact price index for CES with matched model correction in case not all goods are sold in both periods.

Note: All indexes are meant to measure percentage price change between t=0 and t=1. π_i denotes the percentage price change of item i between t=0 and t=1. s_{it} is the expenditure share of good i in period t. C^*_t is the set of items sold in period t. Whenever it is denoted without time index it is assumed that $C^*_0 = C^*_1 = C^*$. $\lambda>0$ is the parameter used in the CES specification of the theoretical model of section 3 and beyond.

Table 2. Actual and estimated levels of CPU price inflation for 10/28/01-03/17/03 for various elasticities of substitution

	With variety bias				Without variety bias			
Elasticity of substitution	1.5	2	4	8	1.5	2	4	8
Theoretical price change	-1%	-16%	-26%	-26%	-31%	-34%	-38%	-40%
Matched model indexes								
Laspeyres	-42%	-37%	-20%	19%	-49%	-46%	-39%	-28%
Paasche	-46%	-46%	-50%	-59%	-52%	-52%	-56%	-62%
Geometric (G0)	-46%	-42%	-26%	11%	-52%	-49%	-42%	-32%
Geometric (G1)	-50%	-50%	-53%	-62%	-55%	-55%	-58%	-64%
Fisher	-44%	-42%	-36%	-30%	-50%	-49%	-48%	-48%
Tornqvist	-48%	-46%	-41%	-35%	-53%	-52%	-51%	-50%
Hedonic price indexes								
Laspeyres	-52%	-46%	-23%	30%	-61%	-59%	-52%	-45%
Paasche	-39%	-44%	-51%	-59%	-41%	-46%	-54%	-59%
Geometric (G0)	-58%	-53%	-31%	17%	-66%	-64%	-56%	-49%
Geometric (G1)	-47%	-50%	-55%	-62%	-49%	-52%	-57%	-61%
Fisher	-46%	-45%	-38%	-27%	-52%	-53%	-53%	-52%
Tornqvist	-53%	-51%	-44%	-33%	-59%	-58%	-57%	-55%



panel (a):
Even though the researcher does not observe the actual quality levels, a_i, the researcher does observe the actual price levels, p_{it}. This panel plots a hypothetical price schedule in two period *t=0,1*.

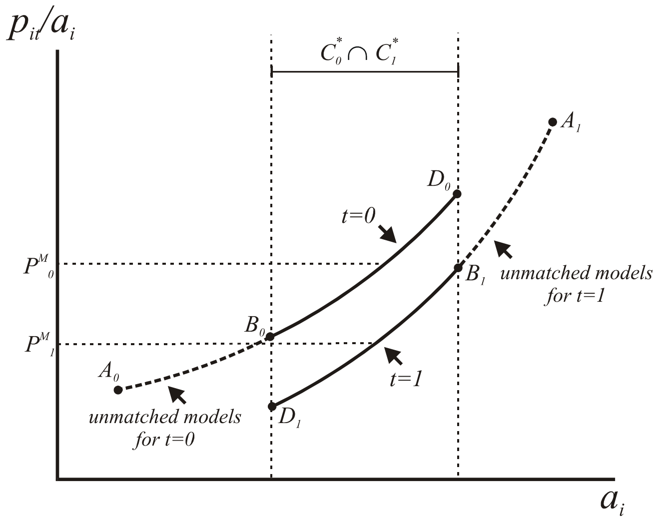
panel (b):
What is relevant for the price aggregate, the percentage change of which is to be measured, is actually the schedule of price per quality unit across models in the market.
This panel plots this schedule. It plots the implied relationship between p_{it}/a_i and a_i.

panel (c):
This panel is identical to panel *(b)*. What is added are:

C^*_t	set of models sold in market at time *t*
P^T_t	theoretical price level at time *t*
B_t-D_t	part of price schedule at time *t* for models that are sold in both periods
A_t-B_t	part of price schedule at time *t* for models that are only sold at time *t*

Note: $P^T_1 > P^T_0$ such that the theoretical inflation level is positive.

Figure 1. Price per quality unit schedule and related notation.

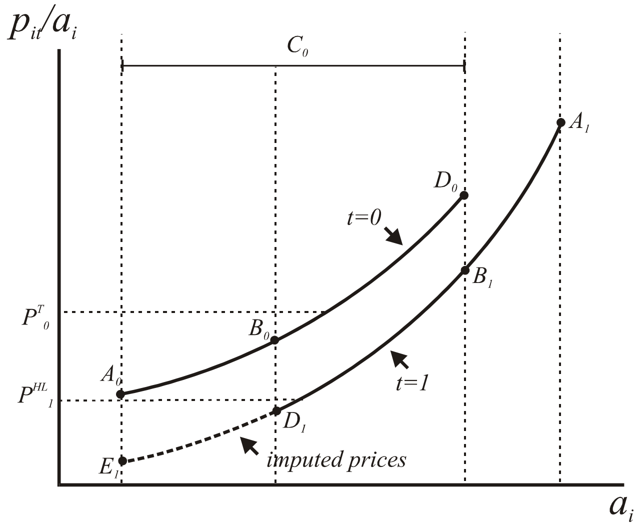


Matched model index: *IP-IQ linking method*
An IP-IQ linked matched model index only compares the prices of models that were sold in both periods and calculates a change in the price index based on these models:

- The set of models sold in both periods is $C^*_1 \cap C^*_2$.
- Calculates price increase between schedules B_0-D_0 and B_1-D_1.
- P^M_t is price level implied by schedule B_t-D_t.
- Measured inflation is change from P^M_0 to P^M_1.

Note: $P^M_1 < P^M_0$ such that the measured inflation level is negative, while the actual level of inflation is positive.

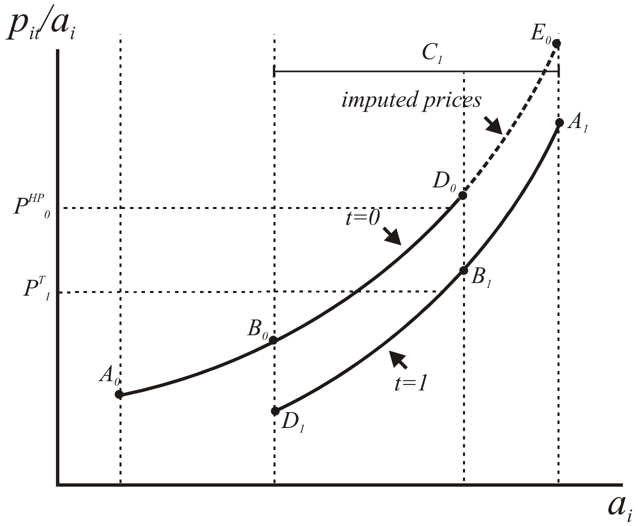
Figure 2. Downward bias in inflation measured using matched model index.



Hedonic index: *Laspeyres index*
A hedonic regression model is used to impute the prices of the unmatched models in period *t=0* for *t=1* in order to calculate a Laspeyres index with *t=0* as baseperiod.

- A Laspeyres index cannot be calculated because price schedule D_1-E_1 is not observed.
- Unobserved part, D_1-E_1, of period *t=1* price schedule is imputed using hedonic model.
- P^{HL}_1 is price level imputed for schedule B_1-E_1.
- Measured inflation is change from P^T_0 to P^{HL}_1.

Note: $P^{HL}_1 < P^T_0$ such that the measured inflation level is negative, while the actual level of inflation is positive.



Hedonic index: *Paasche index*
A hedonic regression model is used to impute the prices of the unmatched models in period *t=1* for *t=0* in order to calculate a Paasche index with *t=1* as measurement period.

- A Paasche index cannot be calculated because price schedule D_0-E_0 is not observed.
- Unobserved part, D_0-E_0, of period *t=0* price schedule is imputed using hedonic model.
- P^{HP}_0 is price level imputed for schedule B_0-E_0.
- Measured inflation is change from P^{HP}_0 to P^T_1.

Note: $P^T_1 < P^{HP}_0$ such that the measured inflation level is negative, while the actual level of inflation is positive.

Figure 3. Downward bias in inflation measured using hedonic price index methods.

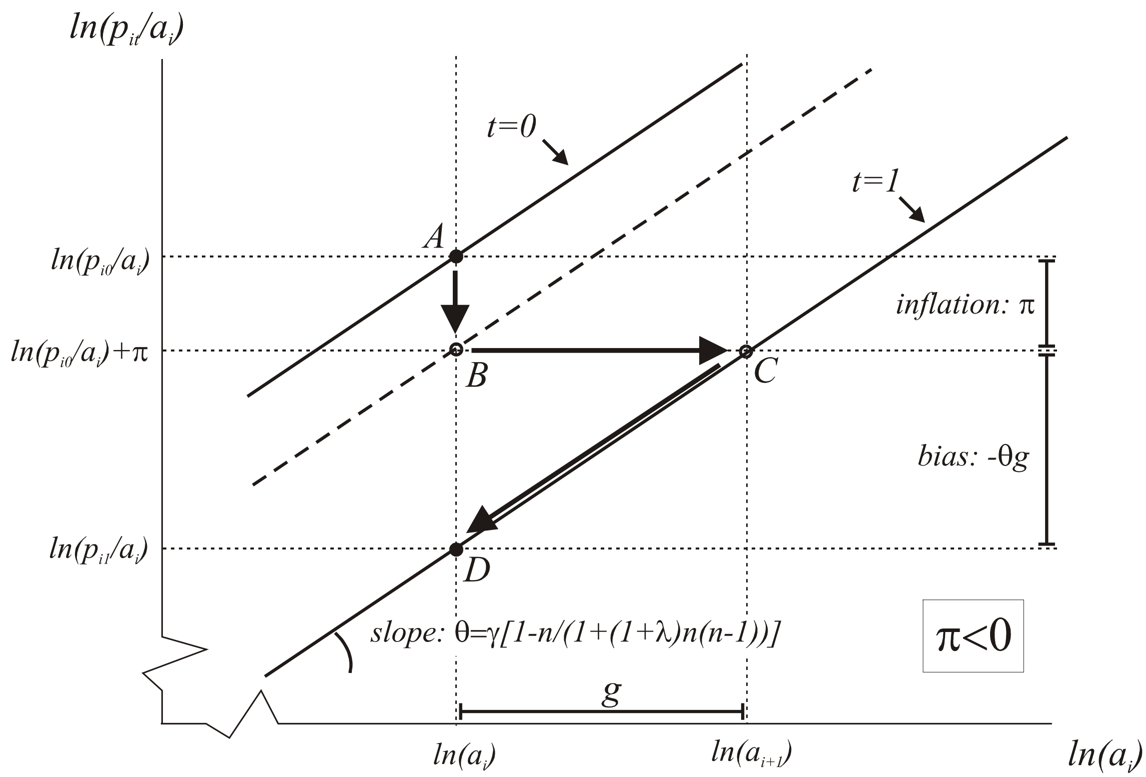


Figure 4. Graphical representation of log-linearization of bias in inflation measures

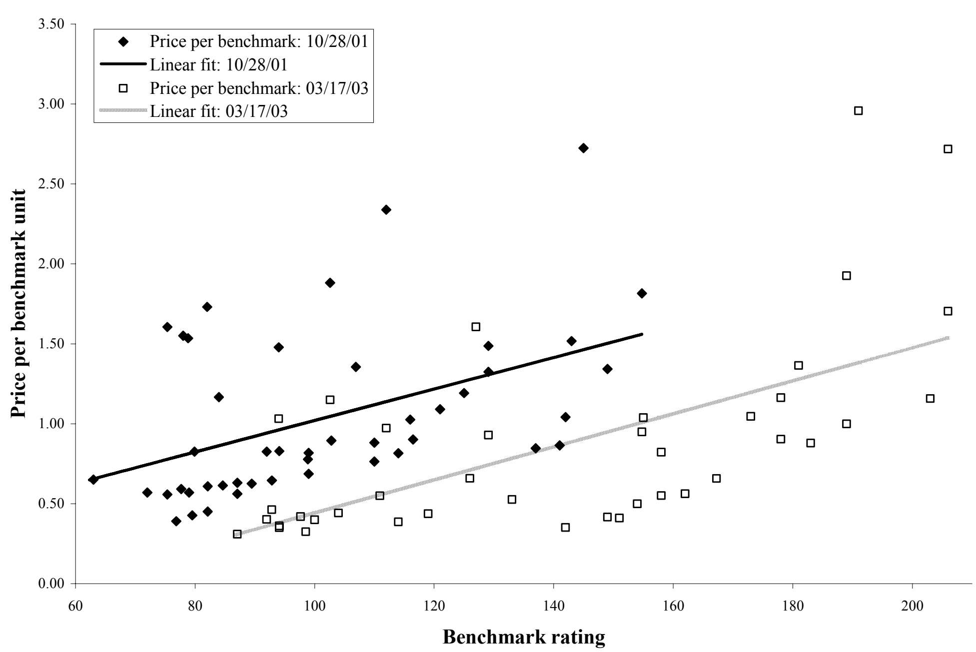


Figure 5. Cross sectional pattern of prices per benchmark unit at beginning and end of sample